

Nama : Hafid ihsan majid

NIM : L200170112

Kelas : D

Modul : 8

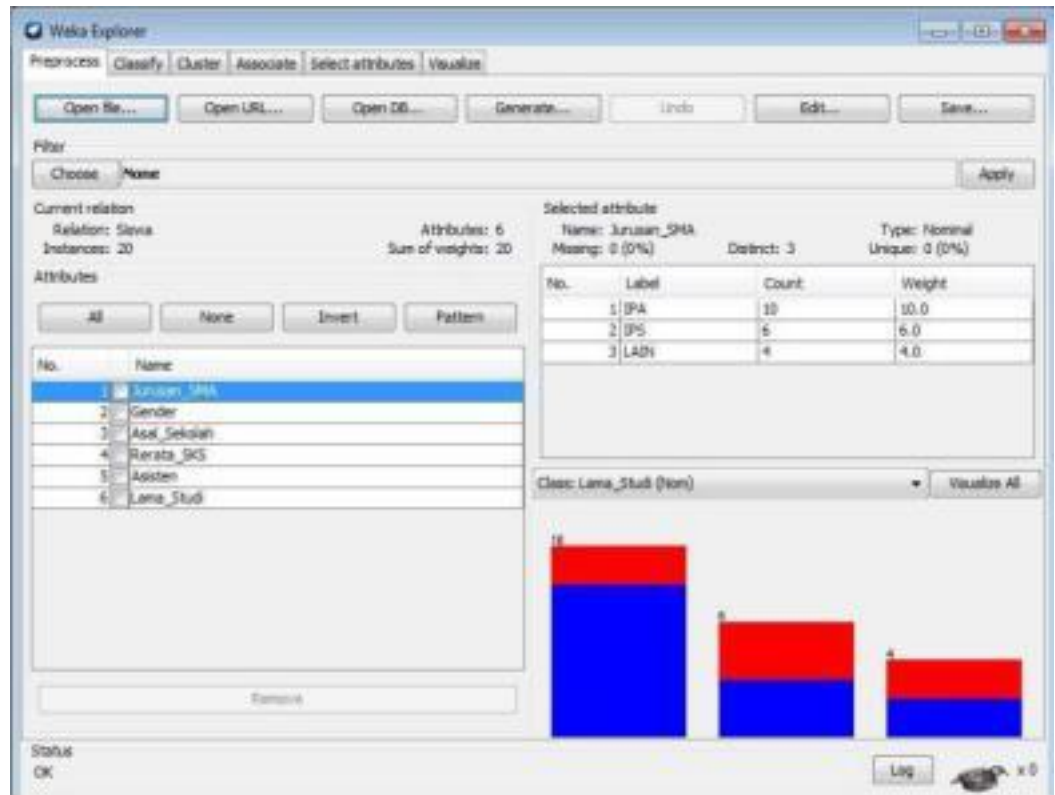
Tugas

1. Berdasarkan tabel berikut, buatlah file dalam format Excel (.xls) dan format ARFF (.arff).
Data ini akan digunakan sebagai **Data Testing**.

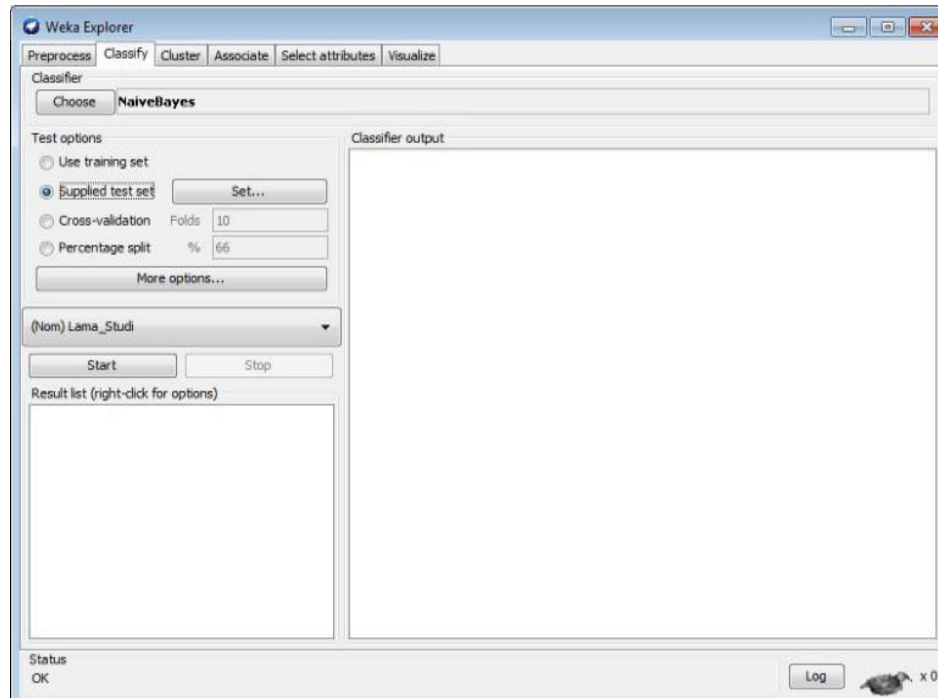
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asiste
2	LAIN	WANITA	SURAKARTA	18	TIDAK
3	IPA	PRIA	SURAKARTA	19	YA
4	LAIN	PRIA	SURAKARTA	19	TIDAK
5	IPS	PRIA	LUAR	17	TIDAK
6	LAIN	WANITA	SURAKARTA	17	TIDAK
7	IPA	WANITA	LUAR	18	YA
8	IPA	PRIA	SURAKARTA	18	TIDAK
9	IPA	PRIA	SURAKARTA	19	TIDAK
10	IPS	PRIA	LUAR	18	TIDAK
11	LAIN	WANITA	SURAKARTA	18	TIDAK

2. Gunakan file ARFF yang dikerjakan pada Tugas nomor 1 dalam Modul 7 sebagai data training. Lakukan prediksi terhadap data testing (ARFF) di atas menggunakan WEKA!

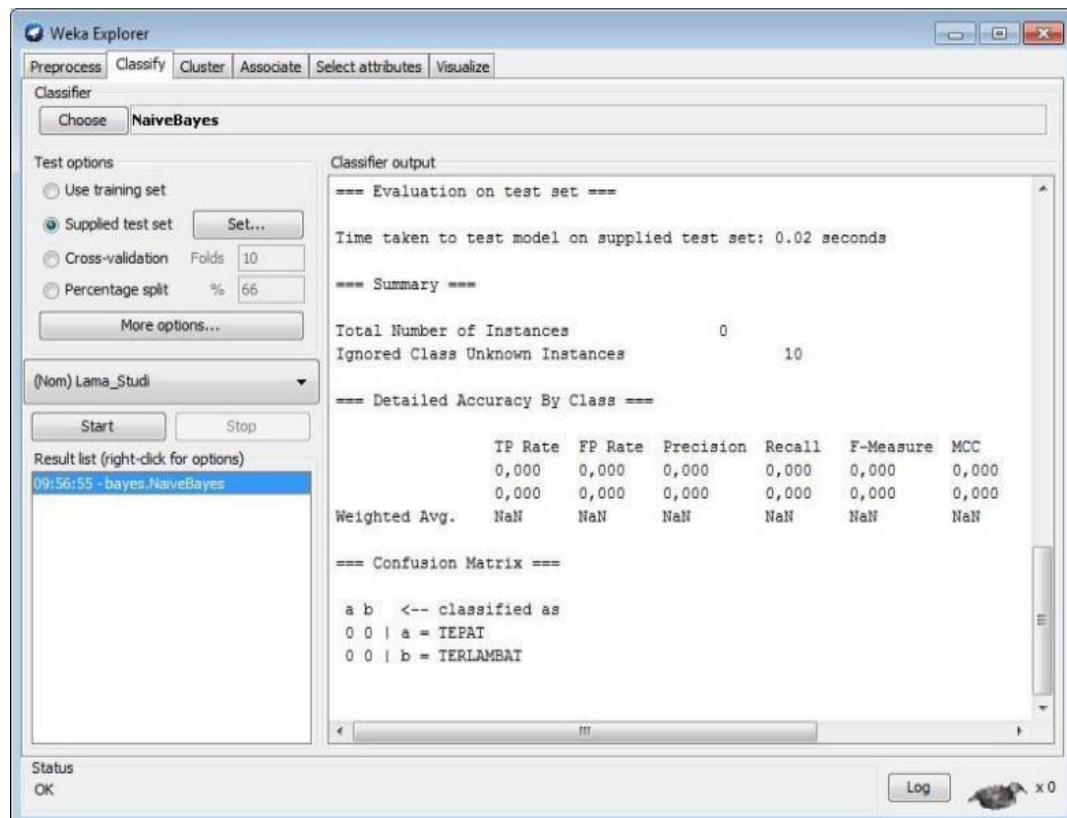
- Membuat File Training



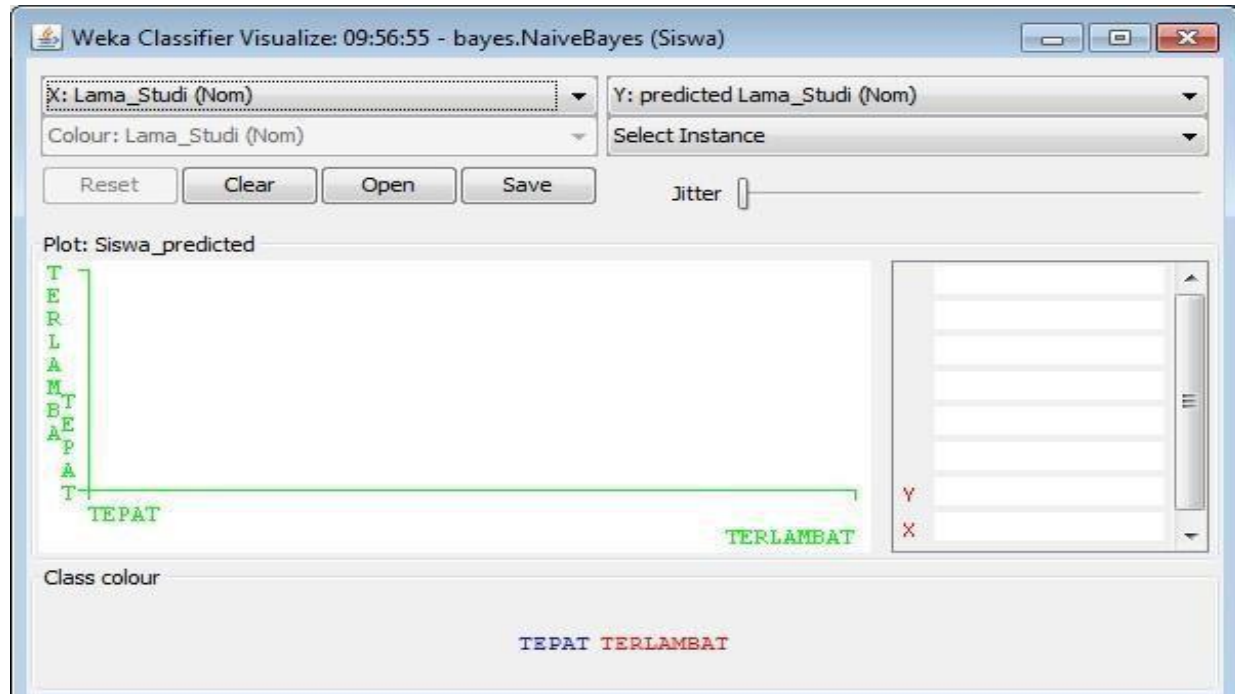
- Memilih Metode Naive Bayes



- Membuka file testing lalu klik kanan pilih visualize classifier errors



- Menyimpan file hasil



- Hasil Prediksi

ARFF-Viewer - D:\dwdm\modul 8\tugas\HasilTugas.arff

File Edit View

HasilTugas.arff

Relation: Siswa_predicted

No.	1: Jurusan_SMA Nominal	2: Gender Nominal	3: Asal_Sekolah Nominal	4: Rerata_SKS Numeric	5: Asisten Nominal	6: prediction margin Numeric	7: predicted Lama_Studi Nominal	8: Lama_Studi Nominal
1	LAIN	WANITA	SURAKARTA	18.0	TIDAK	-0.375862	TERLAMBAT	
2	IPA	PRIA	SURAKARTA	19.0	YA	0.836469	TEPAT	
3	LAIN	PRIA	SURAKARTA	19.0	TIDAK	-0.175169	TERLAMBAT	
4	IPS	PRIA	LUAR	17.0	TIDAK	-0.713206	TERLAMBAT	
5	LAIN	WANITA	SURAKARTA	17.0	TIDAK	-0.546846	TERLAMBAT	
6	IPA	WANITA	LUAR	18.0	YA	0.757815	TEPAT	
7	IPA	PRIA	SURAKARTA	18.0	TIDAK	-0.125076	TERLAMBAT	
8	IPA	PRIA	SURAKARTA	19.0	TIDAK	0.356012	TEPAT	
9	IPS	PRIA	LUAR	18.0	TIDAK	-0.588286	TERLAMBAT	
10	LAIN	WANITA	SURAKARTA	18.0	TIDAK	-0.375862	TERLAMBAT	

- Gunakan file Excel yang dikerjakan pada Tugas nomor 1 dalam Modul 6 sebagai data training. Lakukan prediksi terhadap data testing (Excel) di atas menggunakan RapidMiner!

- Data Training

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sheet1 Cell range: A:F Select All ☒ Define header row: 1

	A	B	C	D	E	F
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi
2	IPS	WANITA	SURAKARTA	18.000	TIDAK	TERLAMBAT
3	IPA	PRIA	SURAKARTA	19.000	YA	TEPAT
4	LAIN	PRIA	SURAKARTA	19.000	TIDAK	TERLAMBAT
5	IPA	PRIA	LUAR	17.000	TIDAK	TERLAMBAT
6	IPA	WANITA	SURAKARTA	17.000	TIDAK	TEPAT
7	IPA	WANITA	LUAR	18.000	YA	TEPAT
8	IPA	PRIA	SURAKARTA	18.000	TIDAK	TERLAMBAT
9	IPA	PRIA	SURAKARTA	19.000	TIDAK	TEPAT
10	IPS	PRIA	LUAR	18.000	TIDAK	TERLAMBAT
11	LAIN	WANITA	SURAKARTA	18.000	TIDAK	TEPAT
12	IPA	WANITA	SURAKARTA	19.000	TIDAK	TEPAT
13	IPS	PRIA	SURAKARTA	20.000	TIDAK	TEPAT
14	IPS	PRIA	SURAKARTA	19.000	TIDAK	TEPAT
15	IPA	PRIA	SURAKARTA	18.000	TIDAK	TERLAMBAT

Previous Next Cancel

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

	Jurusan_SMA * <i>polynominal</i>	Gender * <i>polynominal</i>	Asal_Sekolah * <i>polynominal</i>	Rerata_SKS * <i>integer</i>	Asisten * <i>binominal</i>	Lama_Studi * <i>binominal label</i>
1	IPS	WANITA	SURAKARTA	18	TIDAK	TERLAMBAT
2	IPA	PRIA	SURAKARTA	19	YA	TEPAT
3	LAIN	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT
4	IPA	PRIA	LUAR	17	TIDAK	TERLAMBAT
5	IPA	WANITA	SURAKARTA	17	TIDAK	TEPAT
6	IPA	WANITA	LUAR	18	YA	TEPAT
7	IPA	PRIA	SURAKARTA	18	TIDAK	TERLAMBAT
8	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT
9	IPS	PRIA	LUAR	18	TIDAK	TERLAMBAT
10	LAIN	WANITA	SURAKARTA	18	TIDAK	TEPAT
11	IPA	WANITA	SURAKARTA	19	TIDAK	TEPAT
12	IPS	PRIA	SURAKARTA	20	TIDAK	TEPAT
13	IPS	PRIA	SURAKARTA	19	TIDAK	TEPAT

no problems.

Previous Next Cancel

<new process> - RapidMine Studio Educational 9.3.001 @ asus-PC

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model

Find data, operators, etc. All Studio

ExampleSet (/Local Repository/DataCuaca_Testing) ExampleSet (/Local Repository/DataCuaca_Training) ExampleSet (/Local Repository/DataSekolah_Training) ExampleSet (Apply Model)

Result History

Open in Turbo Prep Auto Model

Filter (20 / 20 examples): all

Row No.	Lama_Studi	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TERLAMBAT	IPS	WANITA	SURAKARTA	18	TIDAK
2	TEPAT	IPA	PRIA	SURAKARTA	19	YA
3	TERLAMBAT	LAIN	PRIA	SURAKARTA	19	TIDAK
4	TERLAMBAT	IPA	PRIA	LUAR	17	TIDAK
5	TEPAT	IPA	WANITA	SURAKARTA	17	TIDAK
6	TEPAT	IPA	WANITA	LUAR	18	YA
7	TERLAMBAT	IPA	PRIA	SURAKARTA	18	TIDAK
8	TEPAT	IPA	PRIA	SURAKARTA	19	TIDAK
9	TERLAMBAT	IPS	PRIA	LUAR	18	TIDAK
10	TEPAT	LAIN	WANITA	SURAKARTA	18	TIDAK
11	TEPAT	IPA	WANITA	SURAKARTA	19	TIDAK
12	TEPAT	IPS	PRIA	SURAKARTA	20	TIDAK
13	TEPAT	IPS	PRIA	SURAKARTA	19	TIDAK
14	TEPAT	IPA	PRIA	SURAKARTA	19	TIDAK

ExampleSet (20 examples, 1 special attribute, 5 regular attributes)

Repository

Import Data

- Samples
- DB (Logos)
- Local Repository (local)
 - Connections (local)
 - data (local)
 - processes (local)
 - DataCuaca_Testing (local - v1, 10/5/19 9)
 - DataCuaca_Training (local - v1, 10/5/19 9)
 - DataSekolah_Training (local - v1, 10/5/19 9)
 - HastiCuaca (local - v1, 10/5/19 10:00 AM)

- Data Testing

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sheet2 Cell range: A:E Select All ☒ Define header row: 1

	A	B	C	D	E
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
2	LAIN	WANITA	SURAKARTA	18.000	TIDAK
3	IPA	PRIA	SURAKARTA	19.000	YA
4	LAIN	PRIA	SURAKARTA	19.000	TIDAK
5	IPS	PRIA	LUAR	17.000	TIDAK
6	LAIN	WANITA	SURAKARTA	17.000	TIDAK
7	IPA	WANITA	LUAR	18.000	YA
8	IPA	PRIA	SURAKARTA	18.000	TIDAK
9	IPA	PRIA	SURAKARTA	19.000	TIDAK
10	IPS	PRIA	LUAR	18.000	TIDAK
11	LAIN	WANITA	SURAKARTA	18.000	TIDAK

← Previous → Next ✗ Cancel

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

	Jurusan_SMA <i>polynomial</i>	Gender <i>polynomial</i>	Asal_Sekolah <i>polynomial</i>	Rerata_SKS <i>integer</i>	Asisten <i>polynomial</i>
1	LAIN	WANITA	SURAKARTA	18	TIDAK
2	IPA	PRIA	SURAKARTA	19	YA
3	LAIN	PRIA	SURAKARTA	19	TIDAK
4	IPS	PRIA	LUAR	17	TIDAK
5	LAIN	WANITA	SURAKARTA	17	TIDAK
6	IPA	WANITA	LUAR	18	YA
7	IPA	PRIA	SURAKARTA	18	TIDAK
8	IPA	PRIA	SURAKARTA	19	TIDAK
9	IPS	PRIA	LUAR	18	TIDAK
10	LAIN	WANITA	SURAKARTA	18	TIDAK

✓ no problems.

← Previous → Next ✗ Cancel

The screenshot shows the RapidMiner Studio interface with the 'Results' view selected. The main window displays a table of 10 examples from the 'ExampleSet (/Local Repository/DataSekolah_Testing)' dataset. The table has 6 columns: Row No., Jurusan_SMA, Gender, Asal_Sekolah, Rerata_SKS, and Asisten. The data is as follows:

Row No.	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	LAJN	WANITA	SURAKARTA	18	TIDAK
2	IPA	PRIA	SURAKARTA	19	YA
3	LAJN	PRIA	SURAKARTA	19	TIDAK
4	IPS	PRIA	LUAR	17	TIDAK
5	LAJN	WANITA	SURAKARTA	17	TIDAK
6	IPA	WANITA	LUAR	18	YA
7	IPA	PRIA	SURAKARTA	18	TIDAK
8	IPA	PRIA	SURAKARTA	19	TIDAK
9	IPS	PRIA	LUAR	18	TIDAK
10	LAJN	WANITA	SURAKARTA	18	TIDAK

Below the table, it states: 'ExampleSet (10 examples, 0 special attributes, 5 regular attributes)'.

- Implementasi Naive Bayes

The screenshot shows the 'Design' view of a RapidMiner workflow. The process is as follows:

- Retrieve DataSekolah...** (Input operator)
- Naive Bayes** (Model training operator)
- Apply Model** (Model application operator)

The workflow is connected in a sequence: Retrieve DataSekolah... → Naive Bayes → Apply Model. The 'Apply Model' operator has two output ports labeled 'res'.

On the right, the 'Parameters' panel for the 'Apply Model' operator is shown, indicating 'No parameters to display.' Below it, there are links for 'Show advanced parameters' and 'Change compatibility (3.3.001)'.

At the bottom, the 'Help' panel for 'Apply Model' is visible, providing a synopsis: 'This Operator applies a model on an ExampleSet.' and a link to 'Jump to Tutorial Process'.

- Hasil

RapidMiner Studio Educational 9.3.001 @ asus-PC

Views: Design Results Turbo Prep Auto Model

Repository: Import Data

Local Repository (Local)

Connections (Local)

Data (Local)

Processes (Local)

DataCuaca_Testing (Local - v1, 10/5/17)

DataCuaca_Training (Local - v1, 10/5/17)

DataSekolah_Testing (Local - v1, 10/5/17)

DataSekolah_Training (Local - v1, 10/5/17)

HasilCuaca (Local - v1, 10/5/17 10:00 AM)

Result History: ExampleSet (Apply Model)

Open in: Turbo Prep Auto Model

Filter (10 / 10 examples): all

Row No.	prediction(L...	confidence(L...	confidence(L...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TERLAMBAT	0.648	0.352	LAIN	WANITA	SURAKARTA	18	TIDAK
2	TEPAT	0.005	0.995	IPA	PRIA	SURAKARTA	18	YA
3	TERLAMBAT	0.650	0.350	LAIN	PRIA	SURAKARTA	19	TIDAK
4	TERLAMBAT	0.868	0.132	IPS	PRIA	LUAR	17	TIDAK
5	TERLAMBAT	0.738	0.262	LAIN	WANITA	SURAKARTA	17	TIDAK
6	TEPAT	0.005	0.995	IPA	WANITA	LUAR	18	YA
7	TERLAMBAT	0.547	0.453	IPA	PRIA	SURAKARTA	18	TIDAK
8	TEPAT	0.321	0.679	IPA	PRIA	SURAKARTA	19	TIDAK
9	TERLAMBAT	0.811	0.189	IPS	PRIA	LUAR	18	TIDAK
10	TERLAMBAT	0.648	0.352	LAIN	WANITA	SURAKARTA	18	TIDAK

ExampleSet (10 examples, 3 special attributes, 5 regular attributes)

RapidMiner Studio Educational 9.3.001 @ asus-PC

Views: Design Results Turbo Prep Auto Model

Repository: Import Data

Local Repository (Local)

Connections (Local)

Data (Local)

Processes (Local)

DataCuaca_Testing (Local - v1, 10/5/17)

DataCuaca_Training (Local - v1, 10/5/17)

DataSekolah_Testing (Local - v1, 10/5/17)

DataSekolah_Training (Local - v1, 10/5/17)

HasilCuaca (Local - v1, 10/5/17 10:00 AM)

Result History: ExampleSet (Apply Model)

Filter (8 / 8 attributes): Search for Attributes

Name	Type	Missing	Statistics
prediction(Lama_Studi)	Binomial	0	Least: TERLAMBAT (3) Most: TERLAMBAT (7) Values: TERLAMBAT (7), TERLAMBAT (3)
confidence(TERLAMBAT)	Real	0	Min: 0.005 Max: 0.868 Average: 0.524
confidence(TEPAT)	Real	0	Min: 0.132 Max: 0.995 Average: 0.476
Jurusan_SMA	Polynomial	0	Least: IPS (2) Most: IPA (4) Values: IPA (4), LAIN (4), ... [1 more]
Gender	Polynomial	0	Least: WANITA (4) Most: PRIA (6) Values: PRIA (6), WANITA (4)
Asal_Sekolah	Polynomial	0	Least: LUAR (3) Most: SURAKARTA (7) Values: SURAKARTA (7), LUAR (3)
Rerata_SKS	Integer	0	Min: 17 Max: 19 Average: 18.100
Asisten	Polynomial	0	Least: YA (2) Most: TIDAK (8) Values: TIDAK (8), YA (2)

Showing attributes 1 - 8

Examples: 10 Special Attributes: 3 Regular Attributes: 5

4. Dari hasil percobaan Tugas nomor 3 di atas, berapakah nilai rerata confidence untuk atribut Lama_Studi dengan nilai TEPAT? Berapakah nilai rerata confidence untuk atribut Lama_Studi dengan nilai TERLAMBAT?

Confidence_TERLAMBAT confidence(TERLAMBAT)	Real	0	Min 0.005	Max 0.868	Average 0.524
Confidence_TEPAT confidence(TEPAT)	Real	0	Min 0.132	Max 0.995	Average 0.476

Jadi nilai rerata confidence untuk atribut Lama_Studi dengan nilai TEPAT adalah **0,524** dan nilai rerata confidence untuk atribut Lama_Studi dengan nilai TERLAMBAT adalah **0,476**

5. Dari hasil percobaan Tugas nomor 3 di atas, berapa orang yang akan lulus TEPAT, dan berapa orang yang akan lulus TERLAMBAT?

Prediction prediction(Lama_Studi)	Binominal	0	Least TEPAT (3)	Most TERLAMBAT (7)	Values TERLAMBAT (7), TEPAT (3)
--------------------------------------	-----------	---	--------------------	-----------------------	------------------------------------

Jadi yang lulus dengan tepat sebanyak **3** orang sedangkan yang lulus terlambat sebanyak **7** orang

6. Prekdisikan ketepatan lama studi si Dewi, jika Dewi adalah seorang WANITA yang berasal dari jurusan IPA pada saat SMA, asal sekolah dari LUAR SURAKARTA, mengambil SKS dengan rata-rata sebanyak 18 SKS tiap semester, dan tidak pernah menjadi Asisten selama kuliah.

- Mengimport Data Dewi

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sheet3 Cell range: A:E Select All ☒ Define header row: 1

	A	B	C	D	E
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
2	IPA	WANITA	LUAR SURAKARTA	18.000	TIDAK

Previous Next Cancel

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

	Jurusan_SMA <i>polynomial</i>	Gender <i>polynomial</i>	Asal_Sekolah <i>polynomial</i>	Rerata_SKS <i>integer</i>	Asisten <i>polynomial</i>
1	IPA	WANITA	LUAR SURAKARTA	18	TIDAK

PREVI...

no problems.

PreviousNextCancel

Result History

ExampleSet (/Local Repository/Data/)

Data

Statistics

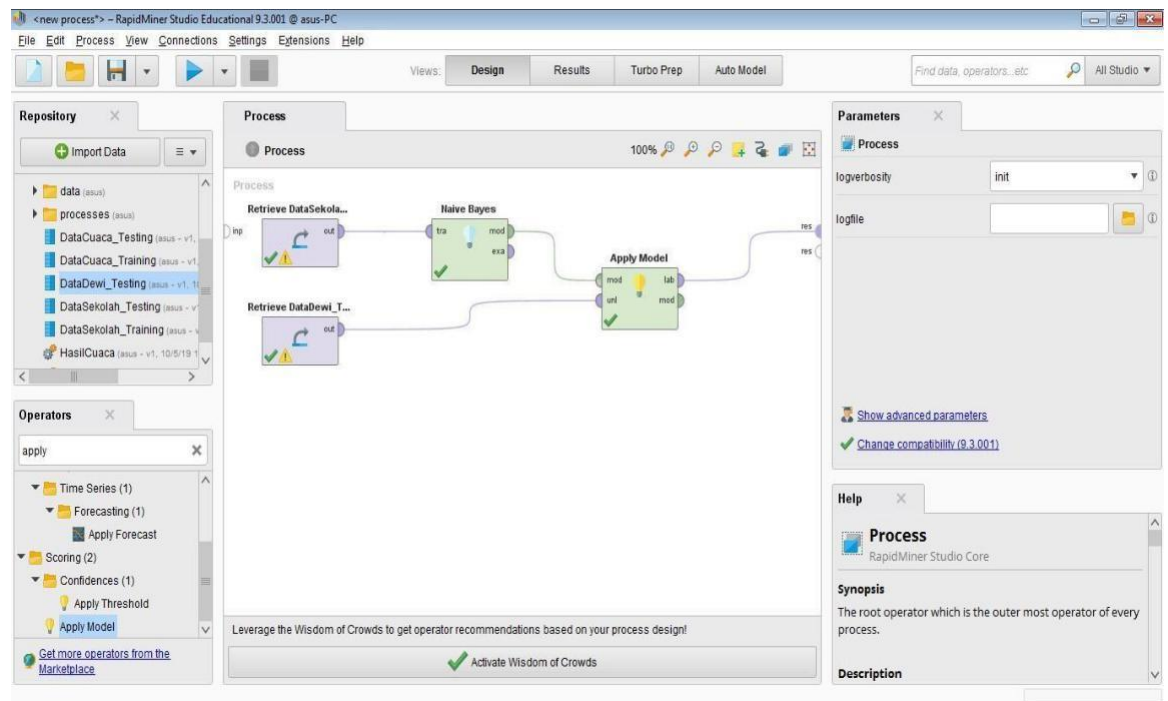
Visualizations

Annotations

Open in Turbo PrepAuto Model

Row No.	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	IPA	WANITA	LUAR SURA...	18	TIDAK

- Mengimplementasikan Naive Bayes



- Hasil

Open in Turbo Prep Auto Model Filter (1 / 1 examples): all

Row No.	prediction(L...	confidence{...	confidence{...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TEPAT	0.457	0.543	IPA	WANITA	LUAR SURA...	18	TIDAK

new process - RapidMiner Studio Educational 9.3.001 @ asus-PC

File Edit Process View Connections Settings Extensions Help

Viewer Design Results **Tools Prep** Auto Model

Find data, operators, etc.

ExampleSet (/Local Repository/DataCuaca_Testing) ExampleSet (/Local Repository/DataCuaca_Training)

ExampleSet (/Local Repository/DataSekolah_Testing) ExampleSet (/Local Repository/DataSekolah_Training)

ExampleSet (/Local Repository/DataDewi_Testing) ExampleSet (/Local Repository/DataSekolah_Testing)

Result History ExampleSet (Apply Model)

Name	Type	Missing	Statistics	Filter (8/8 attributes)	Search for Attributes
✓ Prediction prediction(Lama_Studi)	Binominal	0	Least TERLAMBAT (0)	Most TEPAT (1)	Values TEPAT (1), TERLAMBAT (0)
✓ Confidence_TERLAMBAT confidence(TERLAMBAT)	Real	0	Min 0.457	Max 0.457	Average 0.457
✓ Confidence_TEPAT confidence(TEPAT)	Real	0	Min 0.543	Max 0.543	Average 0.543
✓ Jurusan_SMA	Polynomial	0	Least IPA (1)	Most IPA (1)	Values IPA (1)
✓ Gender	Polynomial	0	Least WANITA (1)	Most WANITA (1)	Values WANITA (1)
✓ Asal_Sekolah	Polynomial	0	Least LUAR SURAKARTA (1)	Most LUAR SURAKARTA (1)	Values LUAR SURAKARTA (1)
✓ Rerata_SKS	Integer	0	Min 18	Max 18	Average 18
✓ Asisten	Polynomial	0	Least TIDAK (1)	Most TIDAK (1)	Values TIDAK (1)

Showing attributes 1 - 8

Examples: 1 Special Attributes: 3 Regular Attributes: 5

Repository

Import Data

Samples

DB (Legacy)

Local Repository

Connections

data (new)

processes

DataCuaca

DataCuaca

DataDewi

DataSekolah

DataSekolah

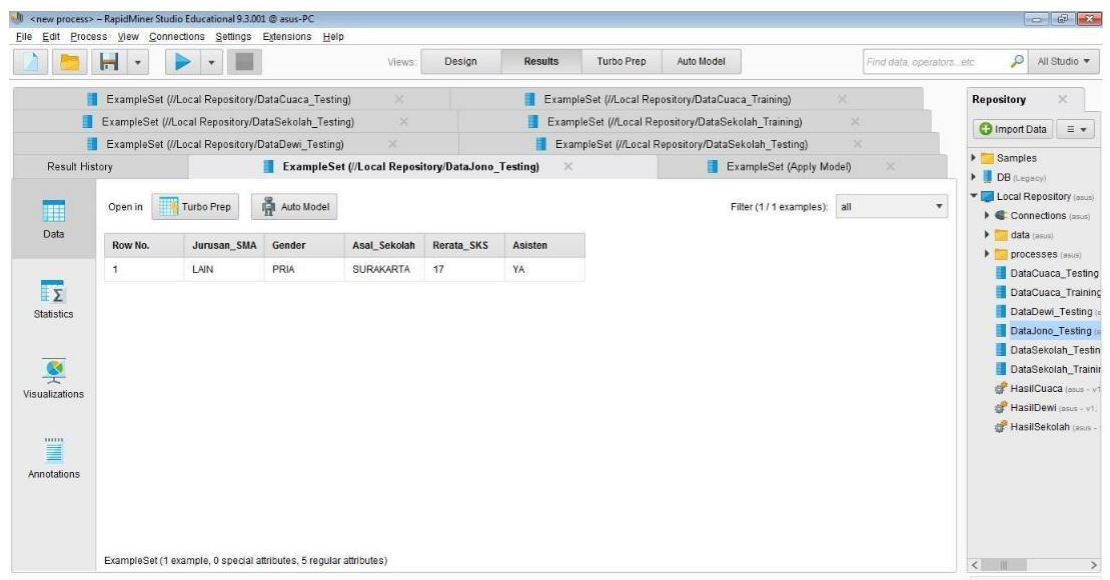
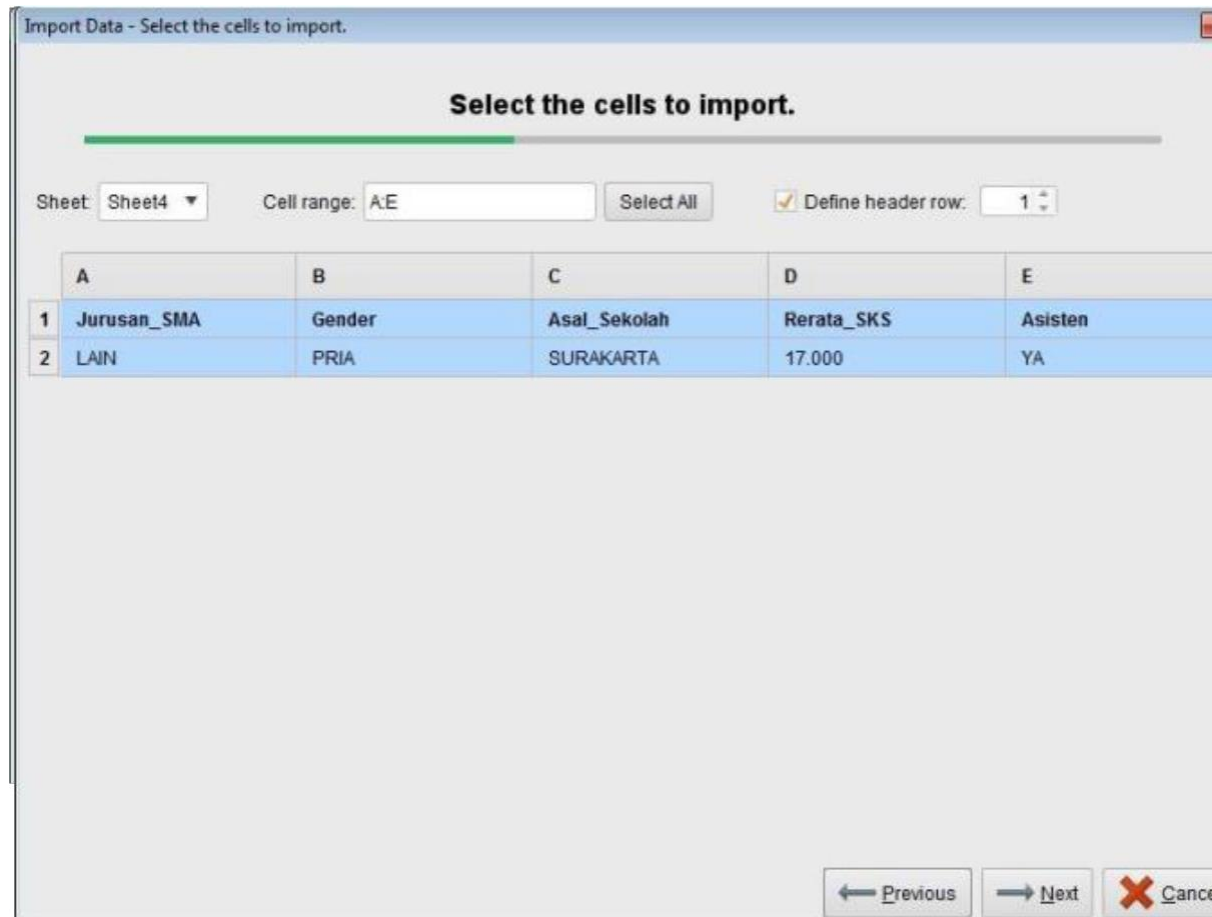
HasilCuaca

HasilSekolah

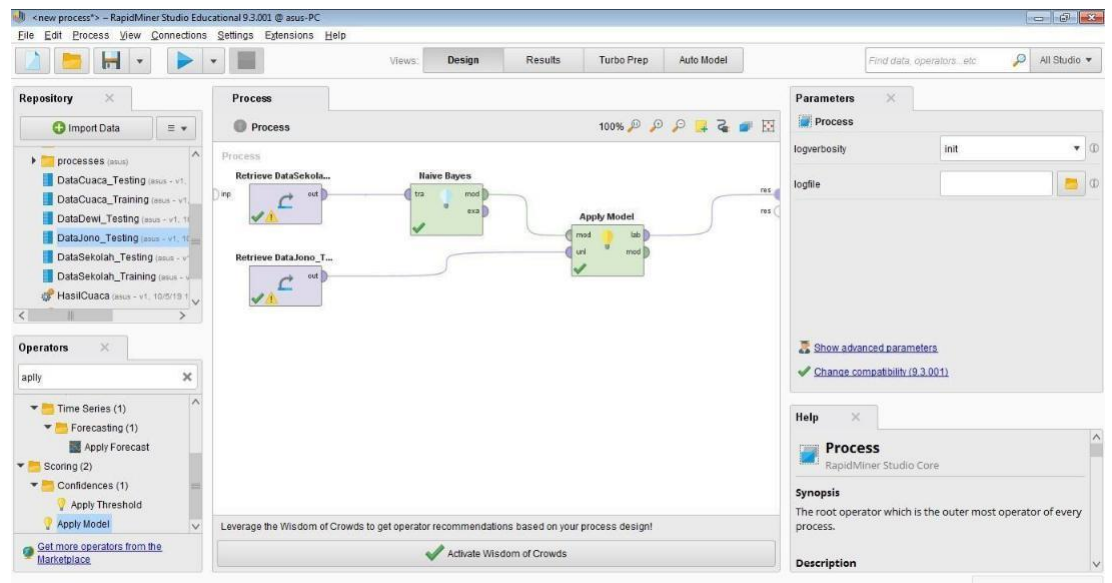
Berdasarkan hasil di atas maka prediksi untuk Dewi dia akan lulus dengan tepat

7. Prekdisikan ketepatan lama studi si Jono, jika Jono adalah seorang PRIA yang berasal dari jurusan IPA pada saat SMA, asal sekolah dari SURAKARTA, mengambil SKS dengan rata-rata sebanyak 17 SKS tiap semester, dan pernah menjadi Asisten selama kuliah.

- Mengimport Data Jono



- Mengimplementasikan Naïve Bayes



- Hasil

Open in [Turbo Prep](#) [Auto Model](#) Filter (1 / 1 examples):

Row No.	prediction(L...	confidence(...	confidence(...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TEPAT	0.076	0.924	LAIN	PRIA	SURAKARTA	17	YA

Result History: [ExampleSet \(Apply Model\)](#)

Name	Type	Missing	Statistics	Filter (8 / 8 attributes):
✓ prediction(Lama_Studi)	Binominal	0	Least TERLAMBAT (0) Most TEPAT (1) Values TEPAT (1), TERLAMBAT (0)	
✓ confidence(TERLAMBAT)	Real	0	Min 0.076 Max 0.076 Average 0.076	
✓ confidence(TEPAT)	Real	0	Min 0.924 Max 0.924 Average 0.924	
✓ Jurusan_SMA	Polynomial	0	Least LAIN (1) Most LAIN (1) Values LAIN (1)	
✓ Gender	Polynomial	0	Least PRIA (1) Most PRIA (1) Values PRIA (1)	
✓ Asal_Sekolah	Polynomial	0	Least SURAKARTA (1) Most SURAKARTA (1) Values SURAKARTA (1)	
✓ Rerata_SKS	Integer	0	Min 17 Max 17 Average 17	
✓ Asisten	Polynomial	0	Least YA (1) Most YA (1) Values YA (1)	

Showing attributes 1 - 8 Examples: 1 Special Attributes: 3 Regular Attributes: 5

Berdasarkan hasil di atas maka prediksi untuk Jono dia akan lulus dengan tepat